Title: Multilocus sequence typing as a tool for studying the molecular epidemiology and population structure of *Brachyspira hyodysenteriae*

Authors: Tom La, Nyree D. Phillips, Belinda L. Harland, Phatthanaphong Wanchanthuek, Matthew I. Bellgard, David J. Hampson

PII: S0378-1135(09)00154-0
DOI: doi:10.1016/j.vetmic.2009.03.025
Reference: VETMIC 4395

To appear in: *VETMIC*

Received date: 27-1-2009
Revised date: 9-3-2009
Accepted date: 12-3-2009

Please cite this article as: La, T., Phillips, N.D., Harland, B.L., Wanchanthuek, P., Bellgard, M.I., Hampson, D.J., Multilocus sequence typing as a tool for studying the molecular epidemiology and population structure of *Brachyspira hyodysenteriae*, *Veterinary Microbiology* (2008), doi:10.1016/j.vetmic.2009.03.025

REVISED

Multilocus sequence typing as a tool for studying the molecular epidemiology and population structure of *Brachyspira hyodysenteriae*

Tom La[1], Nyree D. Phillips[1], Belinda L. Harland[1], Phatthanaphong Wanchanthuek[2], Matthew I. Bellgard[2], David J. Hampson[1]*

Animal Research Institute, School of Veterinary and Biomedical Science, Murdoch University, Murdoch, Western Australia 6150, Australia[1], and Centre for Comparative Genomics, Murdoch University, Murdoch, Western Australia 6150, Australia[2]

Short title: *Brachyspira hyodysenteriae* MLST

*Corresponding author: Tel: +62 89360 2287; fax: +61 89310 41344.

E-mail address: d.hampson@murdoch.edu.au

ABSTRACT

The purpose of this study was to develop and apply a multilocus sequence typing (MLST) scheme to study the molecular epidemiology of *Brachyspira hyodysenteriae*, the aetiological agent of swine dysentery. Sequences of seven conserved genomic loci were examined in 111 *B. hyodysenteriae* strains. Fifty-eight of these previously had been analysed by multilocus enzyme electrophoresis (MLEE), and for some the results of pulsed field gel electrophoresis (PFGE), restriction endonuclease analysis (REA) and/or serotyping also were available. The discriminatory power of these methods was compared. The strains were divided into 67 sequence types (STs) and 46 amino acid types (AATs) by MLST. The Index of Association value was significantly different from zero, indication that the population was clonal. Eleven clonal complexes (Cc) comprising between 2 and 10 STs were recognised. A population snapshot based on AATs placed 77.5% of the isolates from 30 of the AATs into one major cluster. The founder type AAT9 included 13 strains from nine STs that were isolated in Australia, Sweden, Germany and Belgium, including one from a mallard. The MLST results were generally comparable to those produced by MLEE. The MLST system had a similar discriminatory power to PFGE, but was more discriminatory than REA, MLEE or serotyping. MLST data provided evidence for likely transmission of strains between farms, but also for the occurrence of temporal "micro-evolution" of strains on individual farms. Overall, the MLST system proved to be a useful new tool for investigating the molecular epidemiology and diversity of *B. hyodysenteriae*.

Keywords: Brachyspira hyodysenteriae; MLST; swine dysentery; molecular epidemiology; spirochaete

1. Introduction

Swine dysentery (SD) is a mucohaemorrhagic colitis resulting from infection of the porcine caecum and colon with the anaerobic intestinal spirochaete *Brachyspira hyodysenteriae* (Hampson et al., 2006). SD is widespread and problematic in many swine-rearing countries, and the emergence of *B. hyodysenteriae* strains with reduced susceptibilities to various antimicrobials has led to an increased need to understand the molecular epidemiology of the disease (Karlsson et al., 2004; Duinof et al., 2008). To do this, reliable strain typing methods are needed to help trace routes of transmission and monitor reservoirs of *B. hyodysenteriae* infection.

B. hyodysenteriae isolates have been typed using various methods, including serotyping based on reacting extracted lipooligosaccharide with hyperimmune rabbit sera (Baum and Joens, 1979), DNA restriction endonuclease analysis (REA) (Combs et al., 1992; Harel et al., 1994), random amplification of polymorphic DNA (RAPD) (Dougard et al., 1996), DNA restriction fragment polymorphism analysis (RFLP) (Fisher et al., 1997), and pulsed field gel electrophoresis (PFGE) (Atyeo et al., 1999; Fellström et al., 1999). In addition, multilocus enzyme electrophoresis (MLEE) has been used to examine the population structure and diversity of *B. hyodysenteriae* (Lee et al., 1993; Trott et al., 1997). Although MLEE has been useful for differentiating and analysing the relatedness of *B. hyodysenteriae* strains, the technique is slow and cumbersome to perform, and hence it is not suitable for routine use.

Recently, multilocus sequence typing (MLST) has been developed as an alternative method for analysis of bacterial population structure and for discriminating between strains (Maiden et al., 1998; Urwin and Maiden, 2003). The purpose of the present study was to develop a specific MLST system for *B. hyodysenteriae*, based on modifications to the preliminary scheme described for the whole *Brachyspira* genus

(Råsbäck et al., 2007). The method was assessed for its usefulness by applying it to a collection of *B. hyodysenteriae* strains, some of which previously had been analysed by MLEE, PFGE, REA and/or serotyping.

2. Materials and methods

2.1. Brachyspira hyodysenteriae strains

Ninty-five strains of *B. hyodysenteriae* were obtained as frozen stock from the culture collection at the Reference Centre for Intestinal Spirochaetes at Murdoch University. These included 58 representatives of a range of MLEE electrophoretic types previously established in our laboratory (Lee et al., 1993; Trott et al., 1997), as well as 37 other Australian field isolates. The latter included six from feral pigs in Western Australia (WA) (Phillips et al., 2009), and multiple (2-13) recent isolates from each of four large farms, some of which had been tested for their antimicrobial susceptibilities. Many of the older strains that had been analysed by MLEE also had been analysed in our laboratory by PFGE (Atyeo et al., 1999), REA (Combs et al., 1992) and serotyping (Hampson et al., 1989; Combs et al., 1992; Hampson et al., 1997). The 95 strains originated from different States of Australia (n=82), the USA (n=7), Canada (n=3) and the UK (n=3). The names of the strains, their origins and previous results of MLEE, PFGE, REA and serotyping are presented in Table 1. Sequence data for 16 Scandinavian and European strains (AN174/92, AN1082/90, AN3379/98, AN383:2/00, T20, T4, A5677/96, Be45, AN3730/96, AN613/98, AN2420/97, AN360/03, AN551/03, P134/99, E2 and AN1409:2/01) that previously were used in the *Brachyspira* MLST scheme (Råsbäck et al., 2007) were obtained from PubMLST (http://pubmlst.org/) and included in the study. The full collection of

111 strains that was analysed included 101 from commercial pigs, six from feral pigs, two from mallards, one from a rhea and one from a mouse (Table 1).

2.2. Bacterial culture

Spirochaete strains were propagated at 37°C in Kunkle's pre-reduced anaerobic broth containing 2 % (vol/vol) foetal bovine serum and a 1 % (vol/vol) ethanolic cholesterol solution (Kunkle et al., 1986). Cells were harvested from mid-log phase culture by centrifuging at 10,000 x *g*, and counted in a haemocytometer chamber under a phase contrast microscope.

2.3. DNA extraction

The DNeasy kit (QIAGEN Pty Ltd, Doncaster, Australia) was used to extract chromosomal DNA, using the Gram-negative bacterial protocol. Ten ml of a 10^8 cells/ml culture of *B. hyodysenteriae* was centrifuged at 5,000 x *g*. The cell pellet was resuspended in 180 µl of lysis buffer containing 20 µl of proteinase K (10 mg/ml) and incubation at 55 °C for 30 min. After lysis, 180 µl of AL Buffer was added and the sample incubated at 70 °C for 10 min. Two hundred µl of absolute ethanol was added to the sample and this was transferred to a DNeasy spin-column. Ethanol (70%, vol/vol)-based buffers AW1 and AW2 were added sequentially to the columns and centrifuged at 6,000 x *g*. The supernatants were discarded, and the DNA resuspended in sterile water and stored at -20 °C.

2.4. Multilocus sequence typing

Seven of the eight MLST loci previously described for use with the genus *Brachyspira* (Råsbäck et al., 2007) were selected for this study. These were the genes

encoding alcohol dehydrogenase (*adh*), alkaline phosphatase (*alp*), esterase (*est*), glutamate dehydrogenase (*gdh*), glucose kinase (*glpK*), phosphoglucomutase (*pgm*) and acetyl-CoA acetyltransferase (*thi*). The DNA mismatch repair protein gene (*mutS*) was not discriminatory in the previous study, and so it was not used. The primers and PCR conditions were as previously described (Råsbäck et al., 2007), except that only primer pairs ADH-F206 and ADH-R757 were used for *adh*, and ALP-F354 and ALP-R1262 for *alp*.

PCRs were performed in 50 µl reaction mixtures with *Taq* DNA polymerase (Invitrogen, Carlsbad, USA). Each PCR reaction set included *B. hyodysenteriae* strain WA1 as a positive control and double distilled water as a negative control. The PCR conditions were 95 ^{o}C for 2 min, followed by 33 cycles at 95 ^{o}C for 30 sec, 50 ^{o}C for 15 sec, 72 ^{o}C for 1 min, an extension period of 5 min at 72 ^{o}C then cooling to 10 ^{o}C. The PCR products were purified with the AxyPrep^{TM} PCR Clean-up Kit according to the manufacturer's instructions (Axygen Scientific, Inc., Union City, USA).

The purified PCR products were sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, USA) according to the manufacturer's instructions, using the same primers. Sequencing was performed using the 373A sequencing system (Applied Biosystems). Sequence results were analysed and assembled using the ContigExpress component of VectorNTI Advanced 10 (Invitrogen).

To ensure that the sense strand was being used for the analysis, the sequences for each locus were aligned with the original *B. hyodysenteriae* strain WA1 sequence (Råsbäck et al., 2007), using ClustalW (from EMBL-EBI, European Bioinformatics Institute [http://www.ebi.ac.uk/clustalw/]). The aligned loci sequences were trimmed using GeneDoc (Nicholas et al., 1997) such that a 345 bp *adh* allele (nucleotide 303 to

647), 648 bp *alp* allele (nucleotide 460 to 1107), 503 bp *est* allele (nucleotide 288 to 790), 415 bp *gdh* allele (nucleotide 634 to 1049), 687 bp *glpK* allele (nucleotide 320 to 1006), 744 bp *pgm* allele (nucleotide 297 to 1040), and a 745 bp *thi* allele (nucleotide 224 to 968) would be used for the subsequent MLST analysis. The DNA sequences also were translated to predicted amino acid sequences using BioEdit version 7.0.9.0 (Hall, 1999).

2.5. Data analysis

The aligned sequences for each of the seven loci were analysed using the non-redundant databases (NRDB) program (http://pubmlst.org/analysis/) to identify strain sequences that were identical. Each unique nucleotide sequence was assigned a unique allele number. The allelic profile for each isolate was determined and consisted of a line listing the allele number for each locus in turn. Isolates were assigned a sequence type (ST) according to their allelic profiles. Isolates were considered genetically identical and hence of the same ST if they were identical at all seven loci. An analysis was also undertaken with five loci, omitting *adh* and *pgm.* The sequences of the *B. hyodysenteriae* MLST alleles for each locus were deposited at the PubMLST site at Oxford University (http://pubmlst.org/). Unique amino acid types (AATs) predicted from the sequences also were recorded. An MLST dendrogram was constructed from the data matrix of allelic mismatches using the unweighted pair group method with averages method with 1000 bootstrap replicates in START2 (Jolley et al., 2001). Isolates were grouped into clonal complexes (Cc) by the BURST algorithm using the eBURST v3 program (Feil et al., 2004). Within the program a population snapshot was obtained by setting the group definition to 0/7, and assigning a zero for loci without sequence data. The same analysis was used for the AATs. The

degree of linkage disequilibrium in the population was estimated by calculating the Index of Association (I_A) for the 111 strains and the 67 STs (Smith et al., 1993), using the START2 program. A Diversity Index (DI) based on Simpson's index of diversity was calculated for the results of MLST, MLEE, PFGE and REA, as previously described (Hunter and Gaston, 1988).

3. Results

3.1. Sequence types and population structure analysis

The 111 strains analysed were divided into 67 STs and 46 AATs (Table 1). A total of 64 STs and 41 AATs were recognised when *adh* and *pgm* were excluded from the analysis. The corresponding allele numbers assigned for all the STs are shown in the supplementary table, and these raw sequences are recorded in the PubMLST site. Allelic frequency over the seven loci ranged from 9 (for *adh*) to 20 (for *thi*), with a mean of 15.4. Based on the number of strains tested the population had an I_A value of 1.05, whilst based on the number of STs the I_A was 0.175. Significant linkage disequilibrium ($P = 0.000$) was found in both analyses.

[Table 1 about here]

The dendrogram showing the relative relationships of the 67 STs is presented as Figure 1. The basic structure of the tree was unchanged when only five loci were used. STs 1-27, 28-48 and 49-54 each formed a large cluster, with the other groups of STs generally comprising single strains and being arranged in a stepwise fashion with increasing genetic distances between the STs. The two most genetically distinct STs (ST66 and ST67) each contained single strains from Sweden, including one from a mallard duck (Jansson et al., 2004). The small cluster comprising ST61-ST65 contained recent isolates from Western Australia. Overall, the number of isolates in an

ST varied from 1 to 8 (ST19), and the number in an AAT varied from 1 to 14 (AAT17). The 82 strains from Australia were divided into 46 STs and the 29 non-Australian strains were in the other 21 STs. Only in the case of ST52 did strains from different countries share an ST (Germany and Belgium).

[Fig 1 about here]

Eleven clonal complexes (Cc) of STs were identified by e-Burst analysis (Figure 2), and these are marked on Figure 1 and highlighted in Table 1. The Ccs contained between 2 and 10 STs, and between 2 and 26 strains. In seven cases the Ccs were made up of strains just from Australia (Cc1, Cc31, Cc33, Cc36, Cc39, Cc46, Cc64), and these varied in composition from all the strains being from the same farm in the same year (Cc1), and hence likely to be epidemiologically linked, to the strains coming from farms in different States and spanning three decades (Cc46). Cc15 contained 25 Australian strains isolated in different States and decades, and one strain isolated from the USA in the 1980s. Cc51 contained three strains from Germany and one from Belgium, all of which had the unusual phenotype of being indole negative, and which also were considered to be related based on their PFGE patterns (Fellström et al., 1999).

[Fig 2 about here]

For each of the four Australian farms where recent multiple *B. hyodysenteriae* isolates were available, some of the isolates belonged to different STs (Table 1). Farm A had isolates in STs 1, 2, 64 and 65; farm B had isolates in STs 61, 62, 63 and 64; farm C had isolates in STs 3, 49 and 64; farm D had isolates in STs 14, 19 and 21, with all the isolates in the later two STs being resistant to tiamulin.

The six recent isolates from feral pigs in WA all had unique STs. FP1 and FP5 belonged to the same clonal complex (Cc39), together with a WA isolate from a

farmed pig from the 1980s, whilst FP3 was part of a clonal complex (Cc46) with a WA farmed pig isolate from the 1980s, and a Queensland isolate from the 1990s.

A population snapshot obtained by using AATs rather than STs is shown in Figure 3. Thirty (65%) of the AATs were contained in one major cluster, which was made up of two linked sub-clusters. The larger sub-cluster had AAT9 as the founder member, and the other sub-cluster had AAT17 at its centre. This major cluster contained 86 of the strains that were analysed (77.5%). There were three other clusters containing two AATs, one containing four AATs, and each of the other 6 AATs were separate. The founder AAT9 profile was shared by 13 strains from nine STs, comprising seven strains from Australia, three from Sweden (including a strain from a mallard), two from Germany and one from Belgium.

[Fig 3 about here]

3.2. Strain discrimination

Results of the diversity index (DI) using MLST, and a comparison of results for 36 isolates using the different typing methods are summarised in Table 2. MLST was the most discriminatory method of strain typing, with PFGE being the next most discriminatory method, then REA, then MLEE, and then serotyping. The MLST results appeared to be consistent, since in most cases multiple isolates from the same farm isolated in the same year had the same ST. When only five loci were included in the MLST analysis the DI was 0.972.

[Table 2 about here]

4. Discussion

This is the first study in which an MLST scheme has been used to analyse the molecular epidemiology and population structure of *B. hyodysenteriae.* Eighteen *B. hyodysenteriae* strains were included in our previous MLST analysis of species in the

genus *Brachyspira*, but as that study investigated relationships between species, the sequence data were concatenated, and within-species analysis received little attention. With the current larger data set of 111 strains, and by undertaking a comparison of allelic differences at each locus, a more detailed analysis of the population was possible.

Based on calculated I_A values obtained in a previous study using MLEE, it was deduced that *B. hyodysenteriae* is a recombinant species, with an epidemic population structure (Trott et al., 1997). Such populations are basically recombinant, with frequent genetic exchange, but also have clonal groups that have emerged due to them possessing some selective advantage. In the current work, the I_A value obtained was significantly different from zero whether or not it was calculated on the number of strains or number of STs, and this result was consistent with the *B. hyodysenteriae* population being clonal (Smith et al., 1993). The reasons for these different findings with MLEE and MLST are uncertain, but the earlier MLEE study analysed more strains (231) than this MLST study. The current study also was somewhat biased by the fact that 82 (74%) of the strains examined were from Australia, and these included a number of sets of isolates from the same farms. On the other hand, the isolates originated from the 1970s to the 2000s, so a broad timeframe was examined. In the future a more definitive conclusion about the population structure of *B. hyodysenteriae* should become available once MLST results for more strains from around the world are added to the PubMLST database, and the data are reanalysed.

The dendrogram shown as Figure 1 confirmed previous observations that the species is diverse (Lee at al., 1993; Trott et al., 1997), and it helped to depict the relationships between the identified STs. The three large groups of STs (ST1-ST27; ST28-ST48; ST49-ST54) each contained strains isolated in different decades, in

different Australian States and in different countries. Other more distantly related members of the species at the periphery of the dendrogram also were identified. Despite evidence for a general persistence and stability of strains in the larger ST groups, on individual farms there was evidence for the emergence of variants of original strains. A similar phenomenon previously has been observed where PFGE was used to analyse strains from Australian farms, and was described as "micro-evolution" (Atyeo et al., 1999). These apparent minor changes were seen on several occasions, where some strains from the same farm had different STs but belonged to the same Cc (eg ST1 and ST2 in Cc1 on farm A; and ST15, ST19 and ST21 in Cc15 on farm D). In the latter case the isolate in ST15 was recovered in the 1990s and was susceptible to tiamulin, whilst those in ST19 and ST21 were isolated in 2004 and 2005, and were resistant to tiamulin. It appeared that during this period the original strain had acquired mutations in the housekeeping genes that were responsible for assigning it to an ST, and during the same period it also had acquired tiamulin resistance. It was known that tiamulin had been used on this farm to help control SD prior to the emergence of resistance, and this use presumably applied pressure for selection of resistant strains. Mutations in the genes for ribosomal protein L3 and/or 23S rRNA are known to confer resistance to tiamulin in *Brachyspira* species (Pringle et al., 2004).

The MLST data also provided useful molecular epidemiological data at a broader level. For example, some strains with the same ST were found on different farms – and this can be used to provide evidence for the likely transmission of strains between the farms. This was the case for strain WA40; it was both tiamulin resistant and shared ST19 with tiamulin resistant strains from a farm in New South Wales that were isolated in the same year. Based on STs, there was no direct evidence for transmission

of strains between feral and farmed pigs in WA, but in some cases strains were similar enough to suspect that such transmission may have occurred in the recent past. The fact that the feral pig strains were different from each other suggests that these animals are exposed to multiple sources of infection. In farms A, B and C where multiple *B. hyodysenteriae* isolates were obtained, there was clear evidence for the presence of between two and three major distinct strains in each farm, over and above any minor variants of a given strain that were present. This is important information as the different strains may vary in their biological properties, including antimicrobial resistance, and this could influence the success of control programmes. Similarly, where different strains co-exist on a farm this increases the opportunity for exchange of genetic information between them, for example via the activity of the prophage-like gene transfer agents (Humphrey et al., 1997; Motro et al., 2009).

Analysis of AATs in a population snapshot revealed that the majority (86%) of strains belonged to one major cluster (Fig. 3). This greater clustering of AATs compared to STs suggests that, against a backdrop of ongoing mutational changes, there is negative selection on changes in nucleotide sequences that result in changes to amino acid sequences that are essential for protein functionality. Based on this consideration, analysis of AATs appeared to be a useful means of deducing putative ancestral relationships between strains. It was interesting that AAT9 at the centre of the main cluster was shared by strains from different countries, and even included a strain from a mallard that had seemed distantly related from most of the other *B. hyodysenteriae* strains on the basis of its position on the ST dendrogram (Fig. 1). It seems likely that AAT9 represents an ancestral type, from which other AATs have evolved.

As a strain typing method, the use of MLST with seven loci was highly discriminatory, and it was only slightly less discriminatory when used with five loci. Currently it is still recommended to use seven loci until more strains have been examined, and the full extent of diversity in the species has been uncovered. When compared over the same 36 strains, MLST with seven loci had a similar DI to PFGE, but was more discriminatory than REA, MLEE or serotyping. The groupings of strains that were obtained with MLST broadly agreed with those previously established by MLEE and/or PFGE, although there were some exceptions. On the other hand, in agreement with previous observations (Lee et al., 1993), there was no clear association between the serotype and the genetic grouping (ST) of the strains. Compared to MLEE the MLST system was quicker and easier to use, while the results should be more easily compared between laboratories than those of PFGE. An MLST database for *B. hyodysenteriae* has been set up at the PubMLST website, and its future use will allow easy comparisons of other strains isolated in different regions in the world, and from different animal species.

Acknowledgements

The authors thank the Australian Cooperative Research Centre for an International Competitive Pork Industry (the Pork CRC) for financial support for this project. Dr Keith Jolley from the University of Oxford kindly helped to set up the *B. hyodysenteriae* database at PubMLST.

Conflicts of interest

There are no conflicts of interest.

350 **References**

351 Atyeo, R.F., Oxberry, S.L., Hampson, D.J., 1999. Analysis of *Serpulina*
352 *hyodysenteriae* strain variation and its molecular epidemiology using pulsed-
353 field gel electrophoresis. Epidemiol. Infect. 123, 133-138.
354 Baum, D.H., Joens, L.A., 1979. Serotypes of beta-haemolytic *Treponema*
355 *hyodysenteriae.* Infect. Immun. 25, 792-796.
356 Combs, B.G., Hampson, D.J., Harders, S.J., 1992. Typing of Australian isolates of
357 *Treponema hyodysenteriae* by serology and by DNA restriction endonuclease
358 analysis. Vet. Microbiol. 31, 273-285.
359 Dugourd, D., Jacques, M., Bigras-Poulin, M., Harel, J., 1996. Characterisation of
360 *Serpulina hyodysenteriae* isolates of serogroups 8 and 9 by random
361 amplification of polymorphic DNA analysis. Vet. Microbiol. 48, 305-314.
362 Duinhof, T.F., Dierikx, C.M., Koene, M.G., van Bergen, M.A., Mevius, D.J.,
363 Veldman, K.T., van Beers-Schreurs, H.M., de Winne, R.T., 2008.
364 Multiresistant *Brachyspira hyodysenteriae* in a Dutch sow herd. Tijdschr
365 Diergeneeskd. 133, 604-608.
366 Feberwee A., Hampson, D.J., Phillips, N.D., La, T., van der Heijden, H.M.,
367 Wellenberg, G.J., Dwars, R.M., Landman, W.J., 2008. Identification of
368 *Brachyspira hyodysenteriae* and other pathogenic *Brachyspira* species in
369 chickens from laying flocks with diarrhea or reduced production or both. J.
370 Clin. Microbiol. 46, 593-600.
371 Feil, E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., Spratt, B.G., 2004. eBURST:
372 inferring patterns of evolutionary descent among clusters of related bacterial
373 genotypes from multilocus sequence typing data. J. Bacteriol. 186, 1518-30.
374 Fellström, C., Karlsson, M., Pettersson, B., Zimmerman, U., Gunnarsson, A., Aspan,

375 A., 1999. Emended descriptions of indole negative and indole positive isolates
376 of *Brachyspira* (*Serpulina*) *hyodysenteriae*. Vet. Microbiol. 70, 225-238.
377 Fisher L.N., Mathiesen, M.R., Duhamel, G.E., 1997. Restriction fragment length
378 polymorphism of the periplasmic flagellar flaA1 gene of *Serpulina* species.
379 Clin. Diagn. Lab. Immunol. 4, 681-686.
380 Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and
381 analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41, 95-98.
382 Hampson, D.J., Mhoma, J.R.L., Combs, B.G., Buddle, J.R., 1989. Proposed revisions
383 to the serological typing systems for *Treponema hyodysenteriae*. Epidemiol.
384 Infect. 102, 75-84.
385 Hampson, D.J., Atyeo, R.F., Combs, B.G., 1997. Swine dysentery. In: Hampson,
386 D.J., Stanton, T.B. (Eds). Intestinal Spirochaetes in Domestic Animals and
387 Humans. CAB International, Wallingford, UK. pp. 175-209.
388 Hampson, D.J., Fellström, C., Thomson, J.R., 2006. Swine dysentery. In: Straw, B.E.,
389 Zimmerman, J.J., D'Allaire, S., Taylor, D.J. (Eds). Diseases of Swine.
390 Blackwell Publishing, Oxford, UK. pp. 785-805.
391 Harel, J., Belanger, M., Forget, C., Jacques, M., 1994. Characterisation of *Serpulina*
392 *hyodysenteriae* isolates of serotype 8 and 9 from Quebec by restriction
393 endonuclease fingerprinting and ribotyping. Can. J. Vet. Res. 58, 302-305.
394 Humphrey, S.B., Stanton, T.B., Jensen, N.S., Zuerner, R.L., 1997. Purification and
395 characterization of VSH-1, a generalized transducing bacteriophage of
396 *Serpulina hyodysenteriae*. J. Bacteriol. 179, 323-329.
397 Hunter, P.R., Gaston, M.A., 19888. Numerical index of the discriminatory ability of
398 typing systems: an application of Simpson's index of diversity. J. Clin.
399 Microbiol. 26, 2465-2466.

400 Jansson D.S., Johansson, K.E., Olofsson, T., Råsbäck, T., Vågsholm, I., Pettersson,
401 B., Gunnarsson, A., Fellström, C., 2004. *Brachyspira hyodysenteriae* and other
402 strongly beta-haemolytic and indole-positive spirochaetes isolated from
403 mallards (*Anas platyrhynchos*). J. Med. Microbiol. 53, 293-300.
404 Jolley, K.A., Feil, E.J., Chan, M.S., Maiden, M.C., 2001. Sequence type analysis and
405 recombinational tests (START). Bioinformatics 17, 1230-1231.

406 Karlsson M., Aspán A., Landén A., Franklin A., 2004. Further characterization of
407 porcine *Brachyspira hyodysenteriae* isolates with decreased susceptibility to
408 tiamulin. J. Med. Microbiol. 53, 281-285.
409 Kunkle R.A., Harris D.L., Kinyon J.M., 1986. Autoclaved liquid medium for
410 propagation of *Treponema hyodysenteriae*. J. Clin. Microbiol. 24, 669-671.
411 Lee, J.I., Hampson, D.J., Combs, B.G., Lymbery, A.J., 1993. Genetic relationships
412 between isolates of *Serpulina* (*Treponema*) *hyodysenteriae*, and comparison of
413 methods for their subspecific differentiation. Vet. Microbiol. 34, 35-46.
414 Maiden, M.C., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang,
415 Q., Zhou, J., Zurth, K., Caugant, D.A., Feavers, I.M., Achtman, M., Spratt,
416 B.G., 1998. Multilocus sequence typing: a portable approach to the
417 identification of clones within populations of pathogenic microorganisms.
418 Proc. Natl. Acad. Sci. USA. 95, 3140-3145.
419 Motro, Y., La, T., Bellgard, M.I., Dunn, D.S., Phillips, N.D., Hampson, D.J., 2009.
420 Prophage-like gene transfer agents in the pathogenic intestinal spirochaetes
421 *Brachyspira hyodysenteriae, Brachyspira intermedia* and *Brachyspira*
422 *pilosicoli*. Vet. Microbiol. 134, 340-345.
423 Nicholas, K.B., Nicholas, H.B., Deerfield, D.W., 1997. GeneDoc: analysis and
424 visualization of genetic variation. EMBNEW.NEWS 4, 14.

425 Phillips, N.D., La, T., Adams, P.J., Harland, B.L., Fenwick, S.G., Hampson, D.J.,
426 2009. Detection of *Lawsonia intracellularis, Brachyspira hyodysenteriae*
427 and/or *Brachyspira pilosicoli* in feral pigs. Vet. Microbiol. 134, 294-298.
428 Pringle, M., Poehlsgaard, J., Vester, B., Long, K.S., 2004. Mutations in ribosomal
429 protein L3 and 23S ribosomal RNA at the peptidyl transferase centre are
430 associated with reduced susceptibility to tiamulin in *Brachyspira* spp. isolates.
431 Mol Microbiol. 54, 1295-306.
432 Råsbäck, T., Johansson, K-E., Jansson, D.S., Fellström, C., Alikhani, Y., La, T.,
433 Dunn, D.S., Hampson, D.J., 2007. Development of a multilocus sequence
434 typing scheme for intestinal spirochaetes of the genus *Brachyspira.*
435 Microbiology 153, 4074-4087.
436 Smith, J.M., Smith, N.H., O'Rourke, M., Spratt, B.G., 1993. How clonal are bacteria?
437 Proc. Natl. Acad. Sci. U.S.A. 90, 4384-4388.
438 Trott, D.J., Oxberry, S.L., Hampson, D.J., 1997. Evidence for *Serpulina*
439 *hyodysenteriae* being recombinant, with an epidemic population structure.
440 Microbiology 143, 3357-3365.
441 Urwin, R., Maiden, M.C., 2003. Multi-locus sequence typing: a tool for global
442 epidemiology. Trends Microbiol. 11, 479-487.
443
444

445 **Figure captions**

446 **Fig. 1.** Dendrogram constructed from combined individual distance matrices of
447 sequences from seven genes *adh*, *alp*, *est*, *gdh*, *glpK*, *pgm* and *thi* in 111 *B.*
448 *hyodysenteriae* isolates divided into 67 sequence types (STs). The ST numbers that
449 formed part of the previous study by Råsbäck et al. (2007) are marked in bold. The
450 branch lines of the STs in the 11 clonal complexes (Cc) that were identified in the
451 current study are similarly outlined in bold. The length of the scale bar represents 1
452 nucleotide substitution in 100 base pairs of the sequenced gene fragment.

453 **Fig. 2.** Population snapshot obtained with eBURST using nucleotide sequences at
454 seven gene loci, with 111 *B. hyodysenteriae* isolates divided in 67 sequence types
455 (ST). Clusters of related isolates and individual unlinked isolates within the
456 population are displayed as a single eBURST diagram using a group definition of 0/7
457 shared alleles. The numbers represent the ST, and the size of the dots represents the
458 relative number of isolates in the ST. The eleven sets of linked isolates correspond to
459 clonal complexes (Cc), with the predicted founders positioned centrally in the cluster
460 line. The names assigned to the Ccs are indicated on Fig. 1.

461 **Fig. 3.** Population snapshot obtained from the 46 amino acid types (AATs) of 111 *B.*
462 *hyodysenteriae* isolates using eBURST. The parameters used to generate the snapshot
463 were the same as for the data in Fig. 2, except that translated allele sequences were
464 used. There is one major cluster of AATs, made up of two sub-clusters, with the
465 predicted founder (AAT9) positioned centrally in the larger sub-cluster. There are
466 four other linked groups made up of 2 or 3 AATs, and the other six AATs are
467 unlinked. The numbers represent the AAT and the size of the dots represents the
468 relative number of isolates in the AAT.

469

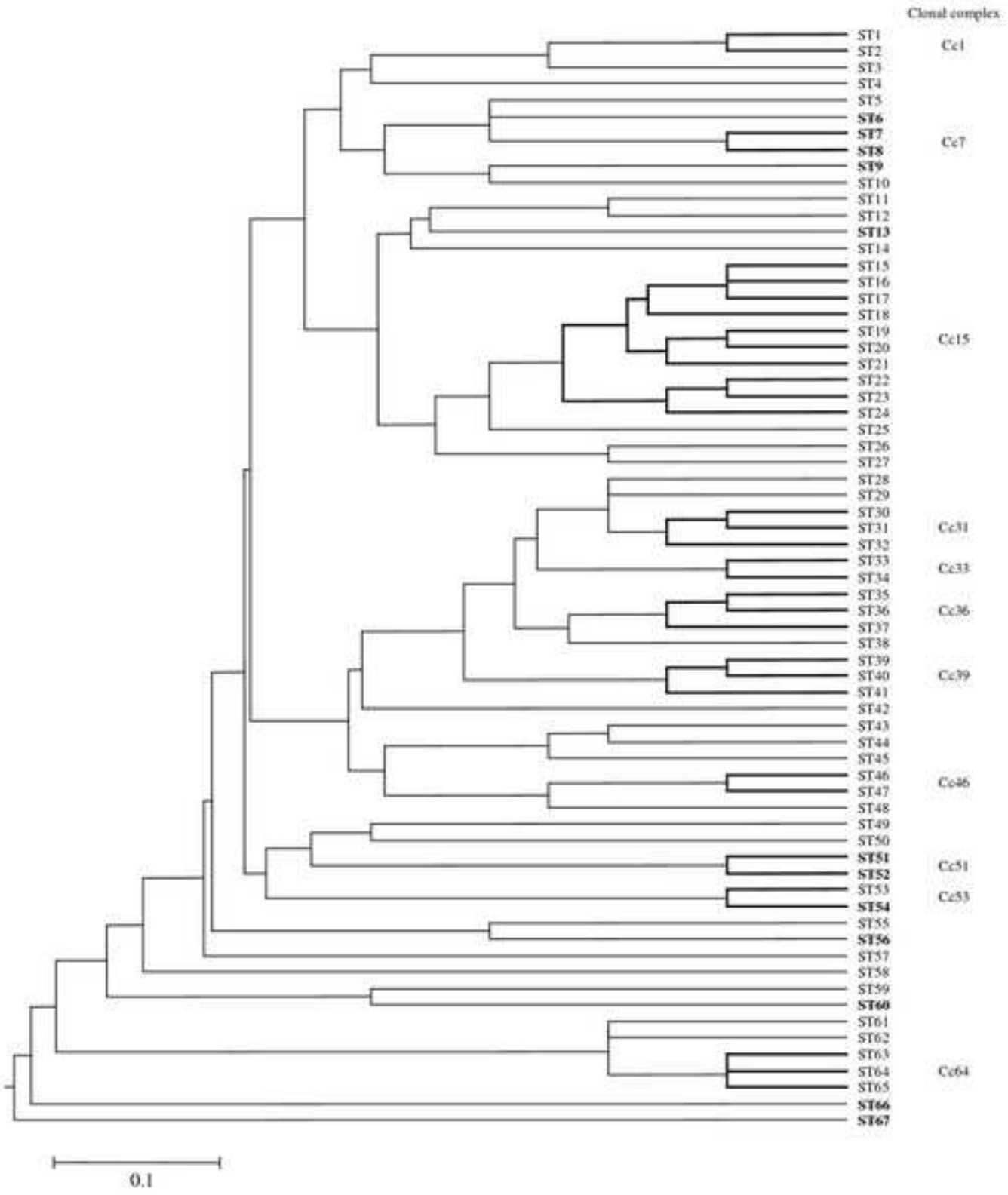

Figure 2

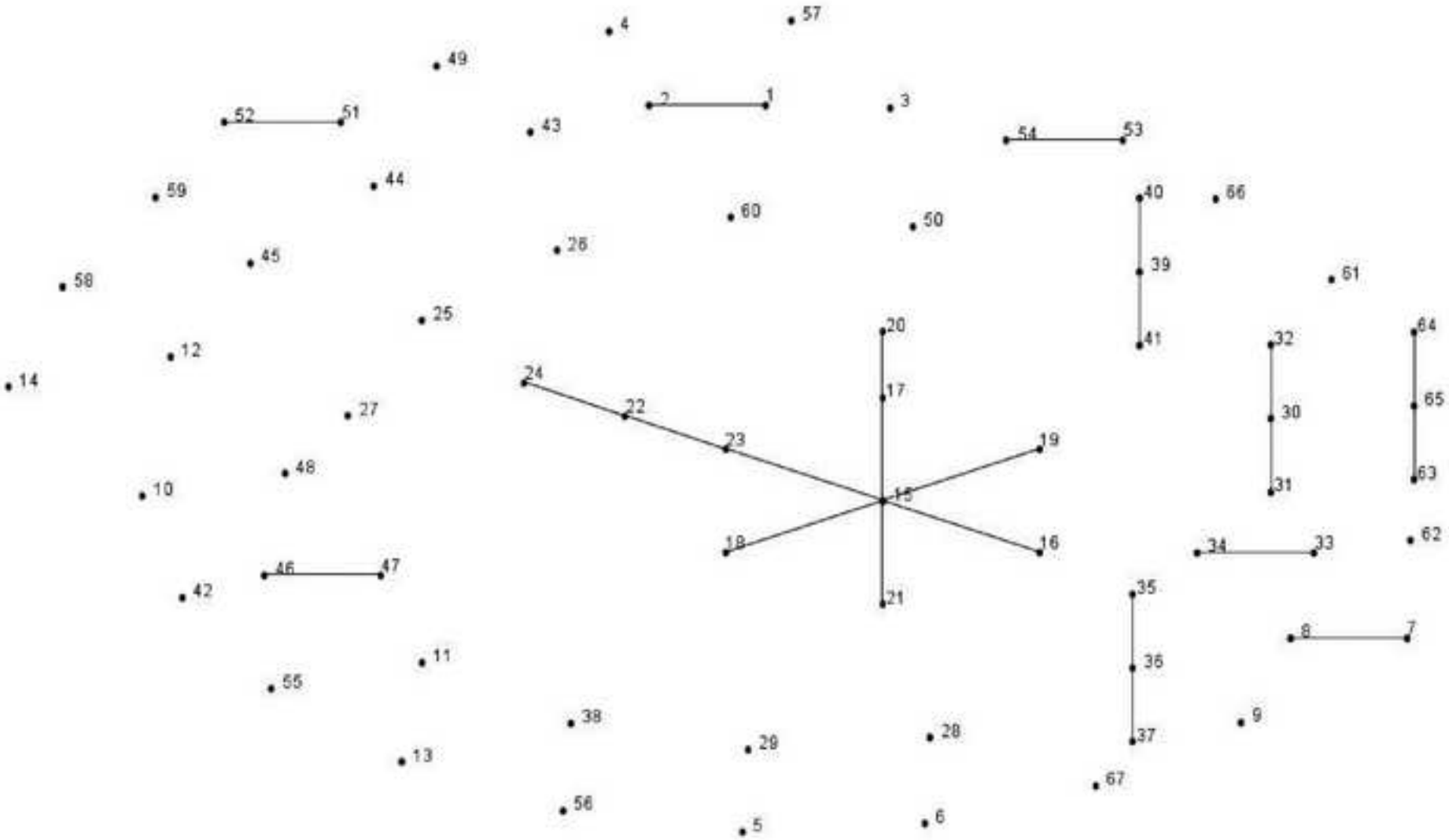
4
57
49
2
1
3
52
51
43
54
53
59
44
40
66
60
50
26
39
61
45
58
25
20
64
12
41
32
24
17
65
14
22
30
27
23
19
48
31
63
10
15
18
16
34
33
62
46
47
35
42
8
7
11
36
55
38
9
29
28
37
13
67
56
5
6

Figure 3

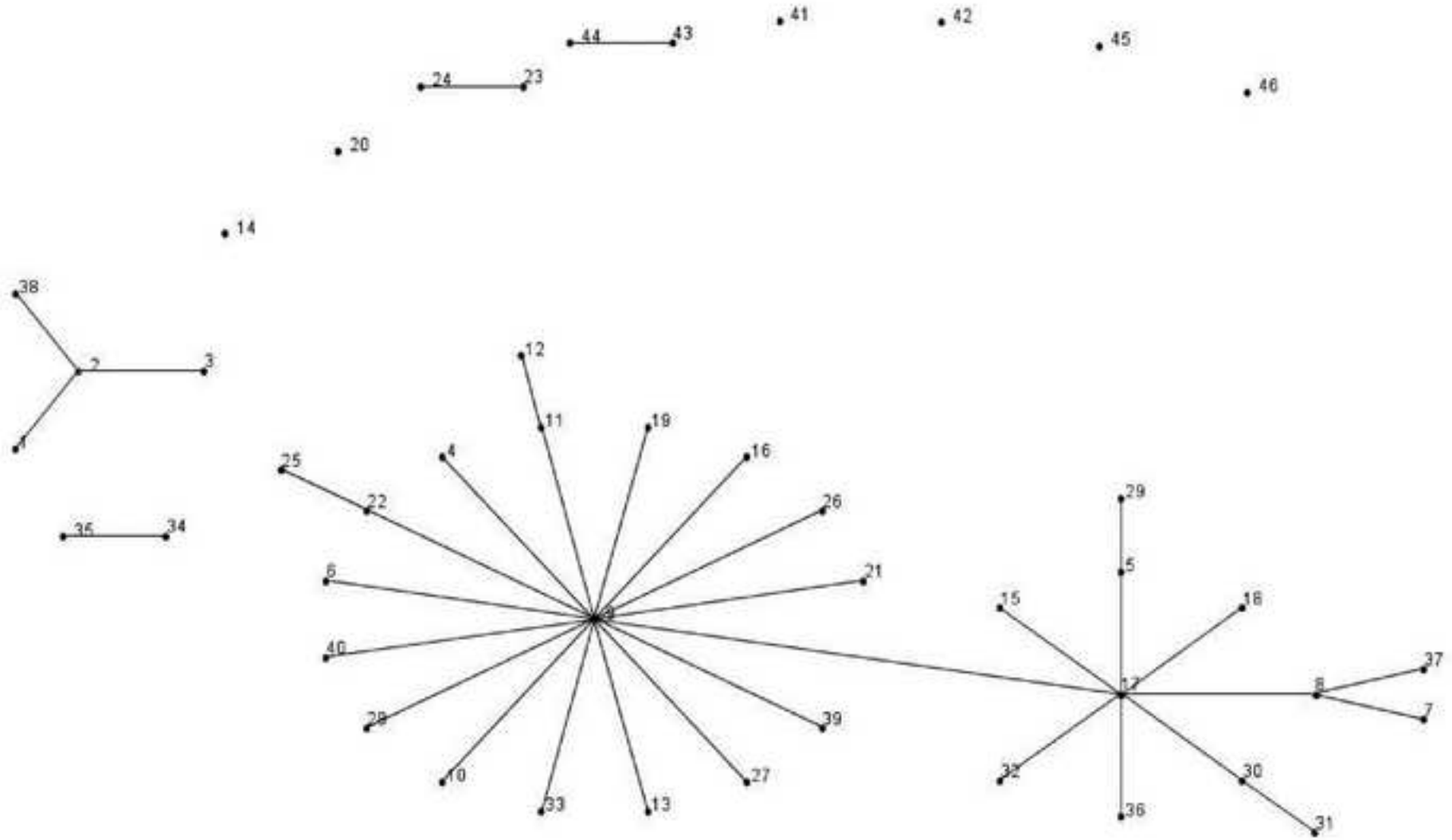
41
42
44
43
45
24
23
46
20
14
38
2
3
12
11
19
4
16
25
22
26
35
34
6
21
40
39
28
10
27
33
13
29
5
15
18
37
12
8
7
32
30
36
31

Table 1

Brachyspira hyodysenteriae strain names and origin, sequence type and amino acid type in MLST, and comparison with previous results of multilocus enzyme electrophoresis, pulsed field gel electrophoresis, DNA restriction endonuclease analysis and serotyping, where these were available.

				Previous typing results			
Strain	Origin[a]	ST[b]	AAT[c]	ET[d]	PFGE[e]	REA[f]	Serogroup[g]
WA46	WA, 2007 (farm A)	1	1	NT	NT	NT	NT
WA41	WA, 2007 (farm A)	2	2	NT	NT	NT	NT
WA52	WA, 2007 (farm A)	2	2	NT	NT	NT	NT
WA53	WA, 2007 (farm A)	2	2	NT	NT	NT	NT
WA56	WA, 2007 (farm A)	2	2	NT	NT	NT	NT
WA57	WA, 2007 (farm A)	2	2	NT	NT	NT	NT
WA42	WA, 2007 (farm A)	2	2	NT	NT	NT	NT
SA12	SA, 2007 (farm C)	3	3	NT	NT	NT	NT
FMV89.3323	Canada, 1989	4	4	12 (9)	NT	NT	K
ACK300/8	USA, 1970s	5	5	20 (15)	NT	NT	B
AN3730/96	Sweden, 1996	6 (6)	6 (6)	NT	NT	NT	NT
AN613/98	Sweden, 1998	6 (6)	6 (6)	NT	NT	NT	NT
P134/99	UK, 1999	7 (9)	7 (9)	NT	NT	NT	NT
E2	UK, unknown	8 (10)	8 (10)	NT	NT	NT	NT
AN360/03	Sweden, 2003	9 (8)	9 (8)	NT	NT	NT	NT
AN551/03	Sweden (mouse), 2003	9 (8)	9 (8)	NT	NT	NT	NT
KF9	UK, 1970s	10	10	22 (12)	NT	G	E
Vic36	VIC, 1991	11	11	6 (4)	B1	H5	B
Vic30	VIC, 1980s	12	12	11 (8)	D1	H8	B
Vic32	VIC, 1980s	12	12	11 (8)	D1	H8	B

AN1082/90	Sweden, 1990	13 (1)	13 (1)	NT	NT	NT	NT
AN174/92	Sweden, 1992	13 (1)	13 (1)	NT	NT	NT	NT
AN3379/98	Sweden, 1998	13 (1)	13 (1)	NT	NT	NT	NT
Vic38	VIC, 1990	14	14	47 (26)	L1	M	B
NSW13	NSW, 1990s	15	15	35 (21)	M2	NT	D
NSW14	NSW, 1990s (farm D)	15	15	35 (21)	M2	NT	D
SA2	SA, 1980s	16	16	36 (22)	K1	L8	D
Vic4	VIC, 1980s	17	15	40 (17)	NT	NT	B
NSW2	NSW, 1990s	18	15	36 (22)	K2	L9	D
NSW42	NSW, 2005 (farm D)*	19	17	NT	NT	NT	NT
NSW45	NSW, 2005 (farm D)*	19	17	NT	NT	NT	NT
NSW46	NSW, 2005 (farm D)*	19	17	NT	NT	NT	NT
NSW49	NSW, 2005 (farm D)*	19	17	NT	NT	NT	NT
NSW27	NSW, 2004 (farm D)*	19	17	NT	NT	NT	NT
NSW28	NSW, 2004 (farm D)*	19	17	NT	NT	NT	NT
NSW32	NSW, 2005 (farm D)*	19	17	NT	NT	NT	NT
WA40	WA, 2005 (farm B)*	19	17	NT	NT	NT	NT
Vic24	VIC, 1988	20	17	43 (18)	H2	L5	D
Vic25	VIC, 1980s	20	17	43 (18)	H2	L5	D
Vic33	VIC, 1980s	20	17	43 (18)	H3	L5	NT
NSW44	NSW, 2005 (farm D)*	21	18	NT	NT	NT	NT
NSW30	NSW, 2005 (farm D)*	21	18	NT	NT	NT	NT
NSW34	NSW, 2005 (farm D)*	21	18	NT	NT	NT	NT
Q1	Q, 1980s	22	16	35 (21)	J1	L2	D
Q21	Q, 1980s	22	16	35 (21)	NT	NT	B
Q3	Q, 1980s	22	16	35 (21)	J1	L2	D
Q8	Q, 1980s	22	16	35 (21)	J2	L3	D

Q9	Q, 1980s	22	16	35 (21)	J2	L3	D
SA1	SA, 1980s	23	16	36 (22)	K1	L8	D
B8044	USA, 1980s	24	19	42	NT	NT	B
SA3	SA, 1980s	25	9	50 (29)	NT	NT	A
B6933	USA, 1980s	26	20	35 (21)	NT	NT	A
Vic23	VIC, 1988	27	9	43 (18)	H1	L5	D
NSW9	NSW, 1991	28	17	NT	M1	NT	NT
Q33a	Q, 1991	28	17	NT	N1	NT	NT
Vic35	VIC, 1980s	29	21	4 (3)	A2	H1	B
Vic2	VIC, 1987	30	9	12 (9)	E2	H6	H
WA26	WA, 1980s	31	9	6 (4)	B2	H5	B
Q10	Q, 1980s	32	22	16 (11)	F1	H9	A
Q11	Q, 1980s	32	22	16 (11)	F1	H9	B
Q14	Q, 1988	32	22	10 (7)	C1	H9	G
Vic31	VIC, 1980s	32	22	10 (7)	C1	H9	B
NSW1	NSW, 1990s	33	23	7 (5)	NT	NT	I
NSW3	NSW, 1990s	34	24	4 (3)	A1	H2	B
Q17	Q, 1990s	35	25	10 (7)	C2	H10	B
Q22	Q, 1990s	36	22	10 (7)	C3	H11	G
WA1	WA, 1980s	36	22	4 (3)	A1	H1	B
WA2	WA, 1980s	36	22	4 (3)	A2	H1	B
WA4	WA, 1980s	36	22	4 (3)	A1	H1	B
WA5	WA, 1980s	36	22	25 (14)	G1	I1	E
WA8	WA, 1980s	36	22	4 (3)	A1	H1	B
WA9	WA, 1980s	36	22	4 (3)	A2	H1	B
WA13	WA, 1980	37	26	4 (3)	NT	NT	B
FP2	WA, 2006	38	22	NT	NT	NT	NT

FP1	WA, 2006	39	27	NT	NT	NT	NT
WA28	WA, 1980s	40	9	16 (11)	F1	H4	A
FP5	WA, 2006	41	27	NT	NT	NT	NT
R301	USA (Rhea), 1996	42	28	20 (15)	NT	NT	A
WA27	WA, 1980s	43	29	38 (19)	I2	J	A
WA14	WA, 1980	44	5	38 (19)	I1	J	A
WA34	WA, 1990s	44	5	25 (14)	P1	NT	E
FP6	WA, 2006	45	17	NT	NT	NT	NT
FP3	WA, 2006	46	30	NT	NT	NT	NT
Q18	Q, 1990s	47	31	10 (7)	C3	H10	G
WA6	WA, 1980s	47	31	25 (14)	G1	I2	E
FP4	WA, 2006	48	32	NT	NT	NT	NT
SA11	SA, 2007 (farm C)	49	9	NT	NT	NT	NT
WA62	WA, 2007 (farm E)	49	9	NT	NT	NT	NT
Q20	Q, 1990s	50	16	49 (28)	NT	NT	A
T4	Germany, 1990s	51 (4)	33 (4)	NT	NT	NT	NT
A5677/96	Germany, 1996	52 (3)	9 (3)	NT	NT	NT	NT
Be45	Belgium, 1990s	52 (3)	9 (3)	NT	NT	NT	NT
T20	Germany, 1990s	52 (3)	9 (3)	NT	NT	NT	NT
B234	USA, 1970s	53	34	1 (1)	NT	A	A
B204	USA, 1970s	54 (12)	35 (12)	20 (15)	NT	C	B
FM88.90	Canada, 1990	55	36	28	NT	NT	J
B78^{T}	USA, 1970s	56 (5)	37 (5)	12 (9)	E1	A	A
P18A	UK, 1970s	57	38	21 (13)	NT	E2	D
VS1	UK, 1970s	58	39	39	NT	NT	NT
B169	Canada, 1970s	59	40	27 (16)	NT	D	C
AN383:2/00	Sweden (mallard),	60 (2)	9 (1)	NT^{h}	NT	NT	NT

	2000						
WA73	WA, 2007 (farm B)	61	41	NT	NT	NT	NT
WA71	WA, 2007 (farm B)	62	42	NT	NT	NT	NT
WA75	WA, 2007 (farm B)	63	43	NT	NT	NT	NT
WA65	WA, 2007 (farm C)	64	44	NT	NT	NT	NT
WA69	WA, 2007 (farm C)	64	44	NT	NT	NT	NT
WA48	WA, 2007 (farm B)	64	44	NT	NT	NT	NT
WA48	WA, 2007 (farm A)	64	44	NT	NT	NT	NT
WA47	WA, 2007 (farm A)	65	43	NT	NT	NT	NT
AN2420/97	Sweden, 1997	66 (7)	45 (7)	NT	NT	NT	NT
AN1409:2/01	Sweden (mallard), 2001	67 (11)	46 (11)	NT	NT	NT	NT

Adjacent strains in the same clonal complex are highlighted with the same background shade. Unshaded strains are not included in a clonal complex.

[a]NSW, New South Wales; WA, Western Australia; VIC, Victoria; Q, Queensland; SA, South Australia. The isolates were recovered from pigs, except where noted in parenthesis. Isolates from feral pigs are marked FP. The year or approximate year of isolation is marked. The farm of origin is marked for recent Australian isolates, if known. Strains that were tested for tiamulin susceptibility and were found to be resistant (MIC ≥ 4 mg/l) are marked with an asterisk (*).

[b]ST, sequence type, with results for the 18 strains previously investigated (Råsbäck et al., 2007) in parenthesis.

[c]AAT, amino acid type, with results for the 18 strains previously investigated (Råsbäck et al., 2007) in parenthesis.

[d]ET, electrophoretic type in multilocus enzyme electrophoresis for 58 strains (Trott et al., 1997), with previous results for 55 strains (Lee et al., 1993) in parenthesis.

[e]PFGE, pulse field gel electrophoresis type for 42 strains (Atyeo et al., 1999). In addition, each of the five pairs of isolates AN3730/96 and AN613/98; AN360/03 and AN551/03; AN1082/90 and AN3379/98; and A5677/96 and Be45 each had identical PFGE patterns (Råsbäck et al., 2007).

[f]REA, restriction endonuclease analysis type for 42 strains (Combs et al., 1992).

[g]Serogroup, for 56 strains (Hampson et al., 1989; Combs et al., 1992; Hampson et al., 1997).

NT, not tested.

Table 2

Diversity Index (DI) for MLST and comparison with other previously described methods of sub-species differentiation of *B. hyodysenteriae* using members of the same subset of 36 strains

Method[a]	Number of strains analysed	Number of types	Frequency (%) of the largest type	Diversity Index (DI)
MLST	111[b]	67	7.2%	0.974
MLST	36[c]	21	19.4%	0.943
PFGE	36	21	11.1%	0.941
REA	36	19	16.7%	0.924
MLEE	36	12	19.4%	0.894
Serotyping	36	6	41.7%	0.719

[a]MLST, multilocus sequence typing (this study)

PFGE, pulse field gel electrophoresis (Atyeo et al., 1999)

REA, restriction endonuclease analysis (Combs et al., 1992).

MLEE, multilocus enzyme electrophoresis (Lee et al., 1993; Trott et al., 1997)

Serogroup, as defined (Hampson et al., 1989; Hampson et al., 1997)

[b]111 refers to the MLST results for all 111 strains in the current study

[c]36 relates to strains that were also used in earlier studies for which there is MLEE, PFGE, REA and serotyping data for all of these strains, as shown in the subsequent rows.

Supplementary table. Sequence types and the corresponding allelic numbers for the seven loci

Sequence type	Loci[a]						
	adh	***alp***	***est***	***gdh***	***glpK***	***pgm***	***thi***
1	2	6	3	1	18	4	3
2	2	6	3	1	17	4	3
3	2	6	3	1	17	10	20
4	2	3	3	1	4	3	16
5	2	10	3	10	8	1	3
6	2	7	3	1	12	1	3
7	2	2	3	12	14	1	3
8	2	2	3	12	11	1	3
9	2	11	3	6	9	1	3
10	2	11	3	1	9	8	13
11	2	11	10	14	3	2	3
12	2	1	10	14	6	2	3
13	2	11	6	6	12	2	3
14	2	14	11	8	3	2	3
15	2	8	3	11	6	2	3
16	2	8	3	5	6	2	3
17	2	8	3	10	6	2	3
18	2	8	3	11	6	2	1
19	2	11	3	11	6	2	3
20	2	11	3	10	6	2	3
21	2	12	3	11	6	2	3
22	2	8	3	4	8	2	3
23	2	8	3	4	6	2	3
24	2	12	3	4	8	2	3
25	2	11	3	9	6	2	17
26	2	16	5	4	6	3	3
27	2	11	3	4	6	3	3
28	2	11	10	10	1	2	1
29	2	11	10	16	1	2	4
30	2	11	10	15	1	2	3
31	2	11	10	7	1	2	3
32	2	11	10	15	1	2	8
33	2	11	10	15	1	9	12
34	2	11	10	15	1	5	12
35	2	11	10	15	13	2	15
36	2	11	10	15	13	2	1
37	2	11	10	18	13	2	1
38	2	11	10	15	7	2	6

39	2	11	9	4	1	2	14
40	2	11	10	4	1	2	14
41	2	11	9	5	1	2	14
42	2	11	8	5	9	2	10
43	2	6	10	10	8	2	5
44	2	6	2	10	8	2	3
45	2	11	2	10	8	2	6
46	2	11	9	10	7	2	6
47	2	11	9	10	7	4	6
48	2	11	10	10	1	4	6
49	2	13	14	17	16	2	6
50	2	8	14	1	13	2	2
51	2	13	4	6	8	2	17
52	2	13	3	6	8	2	17
53	2	6	7	2	4	2	7
54	2	6	7	1	4	2	7
55	2	11	1	10	6	1	9
56	2	2	1	10	5	1	11
57	2	9	5	3	4	4	3
58	4	4	10	4	10	2	2
59	1	16	3	4	2	3	6
60	1	13	3	5	8	1	6
61	8	17	16	8	18	3	14
62	7	17	15	8	18	3	14
63	6	17	11	8	18	3	14
64	2	17	11	8	18	3	14
65	9	17	11	8	18	3	14
66	3	15	12	1	5	6	18
67	5	5	13	13	15	7	19

[a]*adh*, alcohol dehydrogenase; *alp*, alkaline phosphatase; *est*, esterase: *gdh,* glutamate dehydrogenase; *glpK*, glucose kinase; *pgm*, phosphoglucomutasee; *thi*, acetyl-CoA acetyltransferase.